Euro et Math

Robert Rolland

2 Février 2002

1 Euros neufs et vieilles preuves

Le début de l'année 2002 nous a amené nos beaux billets européens tout neufs. Comment résister à la question que tout le monde se posait : les numéros de ces billets contiennent-ils un code détecteur d'erreurs ? Je suis donc allé à la première heure le 1er Janvier en chercher une petite liasse dans le distributeur le plus proche, espérant les trouver, puisque tout neufs, dans l'ordre logique de la numérotation. En effet, voici donc une partie des numéros que j'ai relevés :

U14164027898
U14164027889
U14164027871
U14164027862
U14164027853
U14164027844
U14164027835
U14164027826
U14164027817
U14164027808
U14164027781
U14164027772

On voit tout de suite que **la somme des chiffres modulo** 9 **est constante**. Mais que faire de la lettre ? Hélas tous mes billets commençaient par U. J'entrepris donc de réveiller mes voisins, mes amis qui n'avaient que trop dormi afin d'en trouver un qui aurait un billet préfixé d'une autre lettre. Hélas, personne n'avait le moindre petit billet à me mettre sous la dent. Le croiriez

vous, ils me répondaient tous que disposer d'un billet commençant par un V n'était pas leur souci principal et que d'ailleurs, n'ayant pas complètement cuvé leur alcool du réveillon ils ne s'étaient pas encore procuré les précieuses coupures. En plus ils disaient qu'il était l'heure du petit déjeûner, et que moi-même je ferais aussi bien d'aller me faire cuire un œuf. Heureusement, le site web de la Banque de France donne des photographies de divers billets, on y voit des numéros qui commencent par d'autres lettre que U.

A partir de là une généralisation audacieuse mais raisonnée donne à penser que la règle est la suivante :

On attribue à la lettre le nombre qui donne sa place dans l'alphabet (U est remplacé par 21).**On fait alors la somme modulo** 9 **de ce nombre et de tous les chiffres qui le suivent et on obtient** 8.

Je dois dire que j'étais un peu déçu qu'on nous ait ressorti cette vieille preuve par 9. J'ai donc recherché s'il n'y avait pas un autre mécanisme plus complexe. Je n'en sais rien mais les deux numéros valides

U14164027898
U14164027808

prouvent au moins qu'il existe des cas où une erreur sur un seul chiffre n'est pas détectée.

Que cela soit clair : un code détecteur **n'a pas un rôle de sécurité** destiné à prevenir les contrefaçons. On voit mal un faussaire être assez débile pour fabriquer des billets dont les numéros seraient incohérents. De tels codes sont utilisés par exemple pour détecter des erreurs de saisie.

On peut tout de même se demander si on pouvait faire mieux sans grossir les nombres utilisés.

2 Pouvait-on faire mieux ?

Remarquons qu'on peut considérer que chaque numéro est formé d'une lettre, de 10 chiffres et d'un onzième chiffre pouvant être considéré comme la clé. Cette clé étant calculée de telle sorte que la somme totale (y compris la valeur de la lettre) soit 8 modulo 9. Il y a bien entendu une ambiguité lorsque la clé vaut 9 car alors la clé pourrait être aussi bien 0. Mais je n'ai pas trouvé de billet se terminant par 0.

Ainsi qu'on l'a vu, une telle clé ne permet pas de détecter à coup sûr une unique erreur. On peut citer divers codes pourvus eux aussi d'un seul chiffre de clé et qui eux permettent de détecter à coup sûr une unique erreur.

Exemple 1 : Puique nous sommes dans des histoires bancaires, commençons par la règle de Luhn utilisé sur les numéros de cartes bancaires.
Un numéro de carte bancaire est de la forme

$$a_n a_{n-1} \cdots a_2 a_1 a_0,$$

où les a_i sont des chiffres décimaux qu'on identifiera aux nombres $0, 1, \cdots, 9$.
Sur ces nombres on définit l'application

$$m(x) = 2x \quad si\ 0 \leq 2x \leq 9,$$

$$m(x) = x_1 + x_2 \quad si\ 2x = 10x_1 + x_2,$$

avec $0 \leq x_i \leq 9$. Ainsi $m(x) = 2x \mod 9$ si $0 \leq x \leq 8$ et $m(9) = 9$.

On impose à un numéro de carte bancaire de vérifier (règle de Luhn)

$$a_0 + m(a_1) + a_2 + m(a_3) + \cdots \equiv 0 \quad (10).$$

Le chiffre a_0 peut être considéré comme la clé, calculée en fonction des autres chiffres afin que la règle de Luhn soit vérifiée. Si un des chiffres et un seul est erroné il est facile de voir (en utilisant le fait que m est une bijection) que la cohérence est compromise et l'erreur est détectée. C'est donc déjà mieux. De plus si deux chiffres successifs sont permutés, on s'en aperçoit, sauf si ces deux chiffres sont 0 et 9 (il suffit pour démontrer cela de regarder la fonction $m(x) - x \mod 10$ et de voir que $m(x) - x = m(y) - y$ avec x et y distincts ne peut avoir lieu que si l'un des nombres est 0 et l'autre 9).

Exemple 2 : Le deuxième exemple est le code UPC (Universal Product Code) utilisé pour coder par exemple les produits des supermarchés.
Le code UPC utilise des nombres de 12 chiffres $a_1 \cdots a_{12}$ (11 chiffres pour désigner un produit, et une clé), de telle sorte que

$$\sum_{i=0}^{5} 3a_{2i+1} + \sum_{i=1}^{6} a_{2i}$$

soit divisible par 10.

Si un chiffre du nombre $A = a_1 a_2 \cdots a_{12}$ est modifié, la somme $A_1 = \sum_{i=0}^{5} 3a_{2i+1} + \sum_{i=1}^{6} a_{2i}$ associée à A est modifiée en A'_1 de telle sorte que $|A'_1 - A_1| = a$ ou que $|A'_1 - A_1| = 3a$ avec $1 \leq a \leq 9$. Dans chacun de ces cas $A'_1 - A_1$ n'est pas divisible par 10 et donc A'_1 n'est pas divisible par 10 ce qui permet de détecter l'erreur.

Si le chiffre a_{2i} est permuté avec a_{2i+1}, la somme A_1 est transformée en A'_1 de telle sorte que

$$A'_1 - A_1 = a_{2i+1} - a_{2i} + 3(a_{2i} - a_{2i+1}).$$

Donc

$$A'_1 - A_1 = 2(a_{2i} - a_{2i+1}).$$

En dehors des cas où $|a_{2i} - a_{2i+1}| = 5$, $A'_1 - A_1$ n'est pas divisible par 10 donc l'erreur est détectée.

Hélas il y a là encore des cas particuliers pour lesquels une permutation de deux chiffres consécutifs n'est pas détectée.

Exemple 3 : Donnons un autre exemple instructif, le code ISBN utilisé pour les livres. L'***I**nternational **S**tandard **B**ook **N**umber* utilise des mots de longueurs 10 constitués avec les chiffres $0, 1, ..., 9$ et le symbole X (qui représente le nombre 10) ; le symbole X ne sera utilisé, si nécessaire, que pour la clé.

Exemples : 2 84180 013 X, 2 84225 000 1, 0 471 62187 0, 0 12 163251 2.

Le premier chiffre représente le pays, un bloc de chiffres est attribué à un éditeur, un autre bloc est le numéro donné par l'éditeur, le dernier symbole est la clé, calculée de telle sorte que si $a_1 a_2 ... a_{10}$ désigne un numéro I.S.B.N.

$$\sum_{i=1}^{10} i a_{11-i}$$

soit divisible par 11.

Remarquons que $a_{10} = \sum_{i=1}^{9} i a_i$. Ceci simplifie un peu le calcul de la clé.

Soit A un numéro valide. Appelons A_1 le nombre $\sum_{i=1}^{10} i a_{11-i}$ obtenu à partir des chiffres de A. On sait que A_1 est divisible par 11. Si un chiffre de A est modifié, on obtient alors A' dont le nombre associé A'_1 vérifie

$$|A'_1 - A_1| = ia$$

où $0 \leq i \leq 10$, $1 \leq a \leq 9$. Le nombre ia est premier avec 11, donc A_1' n'est pas divisible par 11, ce qui permet de détecter l'erreur.

Si deux chiffres distincts sont permutés, par exemples ceux d'indices i et j, le nombre A_1 devient A_1' et

$$A_1' - A_1 = i(a_j - a_i) + j(a_i - a_j) = a(i - j),$$

où $a = a_j - a_i$. On a $1 \leq |a| \leq 9$ et $1 \leq |i - j| \leq 9$. Donc $A_1' - A_1$ n'est pas divisible par 11 et A_1' n'est pas divisible par 11.

Cette fois-ci, contrairement aux deux premiers exemples, on a pu détecter outre une erreur sur un chiffre, toutes les permutations de deux chiffres consécutifs. Mais nous avons travaillé modulo 11 ce qui introduit une clé "parasite" X. Bien entendu on pourrait décider de supprimer tous les numéros ayant cette clé parasite. Ce n'est pas commode et fait perdre des représentations valides.

3 Toujours plus fort : la perle rare

Alors la question maintenant est de savoir si on peut faire encore mieux que dans les trois exemples précédents. Autrement dit, peut-on avoir un code avec une clé formé d'un chiffre décimal et qui permette de détecter à coup sûr une unique erreur, ou une permutation de deux chiffres consécutifs.

Notons $a_1 a_2 \cdots a_{i-1} x_i x_{i+1} a_{i+2} \cdots a_n$ un numéro dépendant des deux variables x_i et x_{i+1} où i est une position fixée (autrement dit on regarde ce qu'il se passe lorsqu'on fixe les composantes d'un numéro, sauf deux d'entre elles consécutives. La clé calculée lorsque $x_i = u$ et $x_j = v$ est notée $f(u, v)$ (ici $0 \leq u, v \leq 9$). Si on veut pouvoir détecter à coup sûr un erreur sur un chiffre il est nécessaire que pour v fixée à une valeur v_0, la fonction de u seule, $f(u, v_0)$ soit une bijection de $\{0, \cdots, 9\}$ sur lui même et aussi qu'en fixant u à la valeur u_0 la fonction $f(u_0, v)$ soit bijective. De plus lorsque $u \neq v$ on doit avoir $f(u, v) \neq f(v, u)$. Tout ceci peut s'interpréter de la façon suivante. Construisons une matrice de taille 10×10 pour laquelle les coefficients $A_{u,v}$ sont les $f(u, v)$. Chaque ligne et chaque colonne doit contenir exactement un permutation des 10 chiffres décimaux. Cette matrice est un carré latin. De plus ce carré latin doit être tel que deux éléments qui sont symétriques par rapport à la diagonale principale (sans être sur cette diagonale !) sont toujours distincts.

Il n'y a d'espoir d'existence d'un code vérifiant les propriétés indiquées que si un tel carré latin existe.

Soit D_{10} le groupe diédral d'ordre 10 que nous représenterons comme le groupe des isométries laissant invariant un pentagone régulier. Notons O le centre du pentagone et A_0, A_1, A_2, A_3, A_4 les sommets écrit dans l'ordre (sens trigonométrique). Nous numérotons les transformations de la façon suivante : 0 est l'identité, 1 est la rotation R de centre O et d'angle $\frac{2\pi}{5}$, $2, 3, 4$ sont respectivement les rotations R^2, R^3, R^4. Enfin $5, 6, 7, 8, 9$ sont respectivement les symétries $S_{OA_2}, S_{OA_0}, S_{OA_3}, S_{OA_1}S_{OA_4}$ où S_{OA_i} désigne la symétrie par rapport à l'axe OA_i. On prend comme loi de groupe sur $0, \cdots, 9$ la loi " $*$ " donnée par la composition des transformations. Ainsi $5 * 6 = S_{OA_2} \circ S_{OA_0} = R^4 = 4$ tandis que $6 * 5 = R = 1$.

On introduit aussi la permutation σ de $\{0, \cdots, 9\}$ dans lui même définie par

$$\sigma(0) = 1; \quad \sigma(1) = 5; \quad \sigma(2) = 7; \quad \sigma(3) = 6; \quad \sigma(4) = 2;$$
$$\sigma(5) = 8; \quad \sigma(6) = 3; \quad \sigma(7) = 0; \quad \sigma(8) = 9; \quad \sigma(9) = 4;$$

La matrice A dont les coefficients $A_{u,v}$ vérifient

$$A_{u,v} = u * \sigma(v)$$

est un carré latin ayant la propriété indiquée. Pour voir cela il suffit d'écrire cette matrice.

$$A = \begin{pmatrix} 1 & 5 & 7 & 6 & 2 & 8 & 3 & 0 & 9 & 4 \\ 2 & 6 & 8 & 7 & 3 & 9 & 4 & 1 & 5 & 0 \\ 3 & 7 & 9 & 8 & 4 & 5 & 0 & 2 & 6 & 1 \\ 4 & 8 & 5 & 9 & 0 & 6 & 1 & 3 & 7 & 2 \\ 0 & 9 & 6 & 5 & 1 & 7 & 2 & 4 & 8 & 3 \\ 9 & 0 & 3 & 4 & 8 & 2 & 7 & 5 & 1 & 6 \\ 5 & 1 & 4 & 0 & 9 & 3 & 8 & 6 & 2 & 7 \\ 6 & 2 & 0 & 1 & 5 & 4 & 9 & 7 & 3 & 8 \\ 7 & 3 & 1 & 2 & 6 & 0 & 5 & 8 & 4 & 9 \\ 8 & 4 & 2 & 3 & 7 & 1 & 6 & 9 & 5 & 0 \end{pmatrix}.$$

Remarque : Il est possible de montrer que si on avait pris une opération commutative $\times$ alors quel que soit le choix de la permutation σ la matrice ayant pour coefficients $u \times \sigma(v)$ ne convient pas.

Maintenant, expliquons comment calculer la clé. On calcule a_n en fonction de $a_1, \cdots, a_{n-1}$ de telle sorte que

$$\sigma(a_1) * \sigma^2(a_2) * \cdots * \sigma^{n-1}(a_{n-1}) * a_n = 0.$$

On obtient alors le code convoité.

4 Conclusion

Voilà ce qu'à peu près ils auraient pu construire ...